Problem set 6: Lending and Stablecoins

1 Oracle manipulation

In this problem set, we will examine the cost to manipulate price oracles and the incentives this creates for liquidations of overcollateralized loans. We consider the case where a CFMM with trading function

$$\varphi(R) = \sqrt{R_A R_B}$$

is used as a price oracle for token A (we take token B as the numeraire). For simplicity, we will consider the cost only in terms of the token B tendered to the CFMM. In reality, the attacker also receives some of token A, so the derivations below are overestimates. (Surprisingly, these overestimates are actually asymptotically tight, but we do not show this here.)

a) Assuming the CFMM starts at some price m, derive the cost of manipulating the price oracle to a price $m(1+\epsilon)$. In other words, if $g(\delta)$ is the price of the CFMM after a trade of size δ, compute the smallest δ such that $g(\delta) \geq (1+\epsilon)m = (1+\epsilon)g(0)$. We will denote this cost as $C(\epsilon)$.

Hint. You should use the solutions to the previous problem set.

Note that this assumes that there is always an arbitrageur willing to bring the price back down to m. Since we know transactions are atomic, this can only happen *after* the transaction which moved the price upwards. To deal with this problem, most oracles report a time weighted average price (TWAP) instead of the instantaneous price. We will define the quoted price p as the average price at the end of each of the last T blocks.

b) Assume that the CFMM price is m for blocks $1, \ldots, T-1$. What must the price be at the end of block T be for the oracle to report the price $(1+\epsilon)m$?

c) If the oracle is only manipulated at the end of block T and no point before, what is the cost of manipulating the TWAP price to $(1+\epsilon)m$? Compare this to the total cost of manipulating the quoted price by $1+\epsilon$ in each of the T blocks, $T \cdot C(\epsilon)$, instead. How does each scale in the number of blocks used to average, T?

d) Suggest a more robust statistic that could be used for oracles instead of the mean. Can you find a (reasonable) statistic that makes the single-block attack scale at least linearly in the number of blocks, T?